УДК 519.242:004.42

В.В. ВОЛКОВ

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА И ОБРАБОТКА РЕЗУЛЬТАТОВ ПРИ ПОМОЩИ ПРОГРАММНОГО СРЕДСТВА EOSUPPORT

Описываются возможности программного средства EOSupport, разработанного автором и предназначенного для поддержки принятия решений в задачах, требующих планирования эксперимента и статистической обработки результатов. Приводятся научные результаты некоторых практических применений программы.
Ключевые слова: *планирование эксперимента, дисперсионный анализ, описание поверхности отклика, оптимизация процесса.*

Введение. Существует множество программных средств, предназначенных для автоматизации вычислений и осуществляющих поддержку принятия решений в разнообразных задачах, связанных с планированием эксперимента. Однако практически все они являются разработками зарубежных университетов или фирм, не поддерживают русскоязычный интерфейс и являются платными программными продуктами [4]. Поэтому является актуальной задача создания собственной программы для поддержки планирования экспериментов.

Реализованные методы. Созданная ранее автором функционирующая версия программы EOSupport (Experimental Optimization Support, Система поддержки экспериментальной оптимизации), позволяет строить линейные и квадратичные регрессионные модели (содержащие до 10 факторов) исследуемых процессов, а также проводить статистическую обработку получаемых результатов.

Для построения линейной модели пользователю предоставляется возможность выбора между полным и дробным факторным экспериментом (ПФЭ и ДФЭ соответственно). В случае выбора полного факторного эксперимента в линейную модель можно дополнительно включить произвольные эффекты взаимодействия факторов. При проведении факторного эксперимента используется схема с одинаковым количеством повторных опытов. Эти дополнительные опыты позволяют оценить дисперсию единичного измерения $s^2_{восп}$:

$$s^2_{восп} = \frac{\sum_{i=1}^{N}\sum_{u=1}^{m}\left(y_{iu} - \bar{y}_i\right)^2}{N(m-1)} = \frac{\sum_{i=1}^{N} s_i^2}{N}, \qquad \bar{y}_i = \frac{\sum_{u=1}^{m} y_{iu}}{m}, \tag{1}$$

где $i = 1..N$ - номер строки матрицы планирования; $u = 1..m$ - номер повторного опыта; s_i^2 - дисперсия отклонения повторных опытов в каждой точке плана от среднего значения отклика в этой точке; y - матрица $N \times m$ откликов; число степеней свободы $s^2_{восп}$ равно $f_2 = N(m-1)$.

Коэффициенты регрессионной модели рассчитываются также на основании средних значений $\overline{y}_i$, что обеспечивает высокую точность их оценки:

$$b_j = \frac{\sum_{i=1}^{N} x_{ij}\overline{y}_i}{N}, \; j = 1..l, \tag{2}$$

где l - количество коэффициентов в модели.

Статистический анализ линейной модели состоит в следующем. Сначала проверяется гипотеза об однородности дисперсий s_i^2 по критерию Кохрена. Затем при помощи F-критерия Фишера проверяется гипотеза об адекватности модели и, наконец, для каждого рассчитанного коэффициента b_j можно проверить его значимость по t-критерию Стьюдента [1]. Данный набор критериев позволяет достаточно полно оценить надежность полученных результатов эксперимента.

Пример 1. *Использование ПФЭ для исследования взаимодействий факторов в «игрушечном гольфе».* Рассмотрим эксперимент по забрасыванию мяча на заданное расстояние при помощи специальной установки, называемой «игрушечный гольф» [3]. Управляемые факторы: x_1 - длина биты (или клюшки); x_2 - угол отклонения биты перед ударом; x_3 - масса дополнительного груза биты. В качестве отклика выступает расстояние y, которое преодолеет мяч после удара. Необходимо исследовать взаимодействие между факторами и построить регрессионную модель для предсказания результатов произвольного удара.

Методы испытаний. Для решения поставленной задачи подходит полный факторный эксперимент, позволяющий получить неполную квадратичную модель, содержащую линейные члены и эффекты парных взаимодействий. Интервалы варьирования факторов и некоторая дополнительная информация представлены в табл.1. Сгенерированный EOSupport план эксперимента и результаты непосредственных измерений представлены в табл.2.

Таблица 1

Общие данные об эксперименте

Фактор	Название	Единица	Уровень -1	Уровень +1
x_1	Длина	дюйм	5.50	11.50
x_2	Угол	градус	20.00	60.00
x_3	Масса	кг	0.00	2.00
Отклик	Название	Единица		
y	Расстояние	дюйм		
Тип плана	Полный факторный (варьирование факторов на 2 уровнях)			
Модель	$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$			
Количество повторных опытов	3			

Таблица 2

Матрица планирования и результаты непосредственных измерений

№	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	y_1	y_2	y_3
1	1	1	1	1	1	1	1	116	110	121
2	1	-1	1	1	-1	-1	1	39.0	35.0	37.0
3	1	1	-1	1	-1	1	-1	22.5	24.0	23.0
4	1	-1	-1	1	1	-1	-1	8.00	8.50	8.00
5	1	1	1	-1	1	-1	-1	76.0	89.0	82.0
6	1	-1	1	-1	-1	1	-1	16.0	17.5	16.0
7	1	1	-1	-1	-1	-1	1	16.5	16.0	19.5
8	1	-1	-1	-1	1	1	1	1.75	2.50	2.50

Результаты эксперимента и их обсуждение. В результате опытов были получены коэффициенты уравнения регрессии. Проведен статистический анализ модели в целом и ее коэффициентов в отдельности. Результаты сведены в табл.3.

Таблица 3

Результаты статистического анализа эксперимента

Коэффи-ци-енты	Степень сво-боды	Оценка	t -экспер.	t -крит. (p =0.05)	Вывод
b_0	16	37.8	56.88	2.12	значим
b_1	16	21.8	32.81	2.12	значим
b_2	16	25.0	37.71	2.12	значим
b_3	16	8.16	12.29	2.12	значим
b_{12}	16	14.3	21.48	2.12	значим
b_{13}	16	1.55	2.34	2.12	значим
b_{23}	16	5.22	7.86	2.12	значим
Дисперсия	Ст. свободы	Оценка			
$s^2_{ад}$	1	59.38			
$s^2_{восп}$	16	10.58			
Модель			F -эксп.	F -крит. (p =0.01)	
			5.61	8.5	модель адекватна

Получена следующая регрессионная модель в кодированных единицах:

$$y = 37.8 + 21.8x_1 + 25x_2 + 8.16x_3 + 14.3x_1x_2 +$$
$$+ 1.55x_1x_3 + 5.22x_2x_3,$$

в натуральных единицах:

$$y = 13.47 - 2.76x_1 - 1.03x_2 -$$
$$- 6.68x_3 + 0.24x_1x_2 + 0.52x_1x_3 + 0.26x_2x_3,$$

из которой видно, что из всех парных эффектов взаимодействия наиболее сильным является взаимодействие x_1x_3 (длина биты и масса груза). Все вычисления проводились два раза: первый – при помощи авторской программы EOSupport и второй – при помощи коммерческого продукта Design-Expert trial-версии 7.1.4 фирмы Stat-Ease, Inc. Результаты полностью совпали.

Описание экстремальной области. После получения линейной модели, описывающей локальную область факторного пространства, исследователю надлежит принять решение о дальнейшем ходе исследования; следует ли оптимизировать процесс, двигаясь в направлении градиента (траекторию автоматически рассчитает EOSupport), либо область экстремума уже достигнута и необходимо построение более точной квадратичной модели. Чтобы выяснить это в EOSupport, существует специальная функция оценки кривизны поверхности отклика в области эксперимента. Для этого пользователю предлагается поставить несколько дополнительных опытов в центре плана. Мерой кривизны поверхности отклика может служить разность среднего значения этих откликов и свободного члена линейной модели [2].

Построение квадратичных моделей также легко осуществимо при поддержке EOSupport. Для этого в текущей версии имеется поддержка построения ортогональных центральных композиционных планов (ОЦКП), в которых факторы варьируются на пяти различных уровнях. Достаточно только указать количество факторов, количество повторных экспериментов и количество дополнительных экспериментов в центре плана, и система автоматически рассчитает величину «звездного плеча» α , а также смещение для квадратичных членов. Также можно достроить существующий план ПФЭ или ДФЭ до ОЦКП.

Обычно на практике проведение экспериментов является дорогостоящей и трудоемкой операцией, поэтому при планировании всегда стремятся уменьшить количество точек плана. Чтобы достичь этого без существенных потерь точности, применяется специальная схема, согласно которой повторные опыты, необходимые для оценки дисперсии единичного измерения, ставятся только в центре плана [5]. Использование данной схемы иллюстрирует следующий пример.

Пример 2. *Применение ОЦКП для оптимизации процесса очистки воды.* Вода, полученная после процесса осушения баков для хранения бензина, подвергалась очистке в пробной установке для исследования эффективности совместного влияния коагулянта (сульфата алюминия) и катионового флокулянта (полиэлектролита) группой ученых из Национального технического университета в Афинах [6].

Методы испытаний. Для построения квадратичных регрессионных моделей трех параметров оптимизации (мутность, наличие твердой взвеси, наличие нефтепродуктов) использовалось композиционное планирование. В качестве факторов были выбраны концентрации коагулянта, флокулянта и уровень pH. Автор настоящей статьи использовал полученные исследователями экспериментальные данные при апробации программы EOSupport для проверки вычисленных статистических оценок. Пример ограничивается лишь одним параметром оптимизации – уровнем замутненности воды, подвергшейся очистке. Общие данные представлены в табл.4 и 5.

Таблица 4

Общие данные об эксперименте

Фактор	Название	Единица	-1.414	-1	0	+1	1.414
x_1	Сульфат алюминия	мг/л	4	19	55	91	106
x_2	Полиэлектролит	мг/л	1	5	15	25	29
x_3	pH		5	5.5	7	8.5	9
Отклик	Название	Единица					
y	Мутность	%					
Тип плана	ОЦКП (варьирование факторов на 5 уровнях)						
Модель	$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2$						
Количество опытов в центре плана		4					

Таблица 5

Матрица планирования и результаты непосредственных измерений

№	x_0	x_1	x_2	x_3	y	№	x_0	x_1	x_2	x_3	y
1	1	1	1	1	95.5	10	1	1.41	0	0	56.8
2	1	-1	1	1	69.7	11	1	0	-1.41	0	17.6
3	1	1	-1	1	11.6	12	1	0	1.41	0	41.5
4	1	-1	-1	1	25.6	13	1	0	0	-1.41	35.7
5	1	1	1	-1	68.9	14	1	0	0	1.41	58.8
6	1	-1	1	-1	55.8	15	1	0	0	0	30.9
7	1	1	-1	-1	56.8	16	1	0	0	0	28.7
8	1	-1	-1	-1	17.6	17	1	0	0	0	20.2
9	1	-1.41	0	0	52.6	18	1	0	0	0	25.9

Результаты эксперимента и их обсуждение. После проведения опытов были получены коэффициенты квадратичного уравнения регрессии. Проведен статистический анализ модели. Результаты представлены в табл.6.

Таблица 6

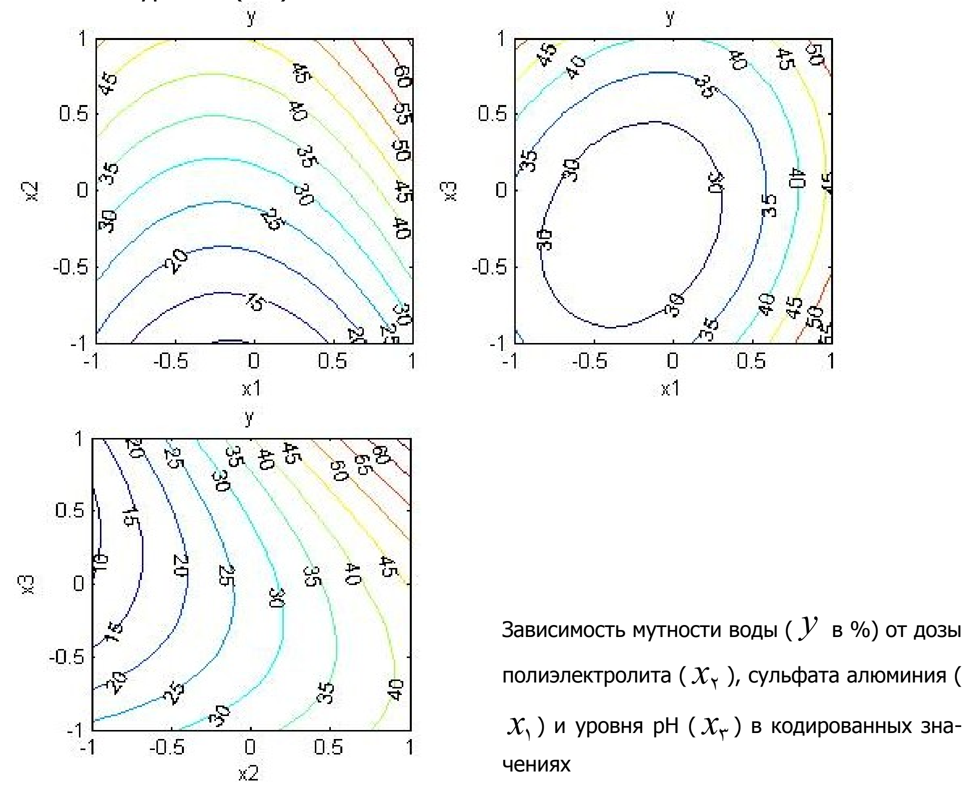
Результаты эксперимента

Коэффи-циент	$b_{٠}$	$b_{١}$	$b_{٢}$	$b_{٣}$	$b_{١٢}$	$b_{١٣}$	$b_{٢٣}$	$b_{١١}$	$b_{٢٢}$	$b_{٣٣}$
Оценка	26.95	5.84	17.68	3.00	1.71	-5.06	9.71	13.35	0.78	9.63
Дисперсия		Степень свободы		Оценка						
$s^{٢}_{ад}$		5		267.45						
$s^{٢}_{восп}$		3		21.41						
Модель		F -эксп.		F -крит. (p =0.05)			Вывод			
		12.49		9.0			модель адекватна			

Особый интерес представляют графики полученной поверхности отклика, представленные на рисунке, определяемой уравнением (в кодированных значениях факторов):

$$y = ٢٦.٩٥ + ٥.٨٤x_{١} + ١٧.٦٨x_{٢} + ٣x_{٣} + ١.٧١x_{١}x_{٢} - ٥.٠٦x_{١}x_{٣} +$$
$$+ ٩.٧١x_{٢}x_{٣} + ١٣.٣٥x_{١}^{٢} + ٠.٧٨x_{٢}^{٢} + ٩.٦٣x_{٣}^{٢}.$$

На каждом из трех графиков отсутствующий фактор зафиксирован на основном уровне (0.0).



Зависимость мутности воды (y в %) от дозы полиэлектролита ($x_{٢}$), сульфата алюминия ($x_{١}$) и уровня pH ($x_{٣}$) в кодированных значениях

Из рисунка хорошо видно, что мутность воды в значительной степени зависит от дозы полиэлектролита и в меньшей степени от уровня pH и дозы сульфата алюминия. Кроме того, мутность имеет минимум около

точки (-۰.۳, ۰, -۰.۲) на уровне около 25%. Результаты, полученные при помощи программы EOSuppport, хорошо согласуются с результатами исследователей, изложенными в [6].

Выводы. Программное средство EOSupport планируется в будущем расширить дополнительными возможностями, такими как поддержка ротатабельных планов, планов Бокса-Бенкена, функциями разбиения плана на блоки, вычисления p-значений для проверки гипотез и др. Наряду с улучшением средств визуализации полученных результатов всё это позволит существенно повысить эффективность применения EOSupppot.

Библиографический список

1. *Адлер Ю.П.* Планирование эксперимента при поиске оптимальных условий / Ю.П. Адлер, Е.В.Маркова, Ю.В.Грановский. - М.: Наука, 1976.

2. *Ахназарова С.Л.* Оптимизация эксперимента в химии и химической технологии. / С.Л. Ахназарова, В.В.Кафаров. - М.: Высшая школа, 1978.

3. Leonard Lye, Teaching DOE with the DOE-Golfer, Faculty of Engineering and Applied Science, Memorial University, St. John’s, Newfoundland, Canada, http://www.statease.com/golftoy.html.

4. Planning for the unexpected, Scientific Computing World: February/ March 2006, http://www.scientific-computing.com.

5. Rajender Parsad, P. K. Batra, Response Surface Designs, Indian Agricultural Statistics Research Institute (I.A.S.R.I.), New Delhi, http://www.iasri.res.in/iasriwebsite/.

6. F. Rigas, P. Panteleos, C. Laoudis, Central Composite Design in a Refinery’s Wastewater Treatment by Air Flotation, National Technical University of Athens, ECOTECH Ltd, Athens, Greece, 2000.

Материал поступил в редакцию 26.02.08.

V. V. VOLKOV

DESIGN OF EXPERIMENTS AND ANALYSIS OF THEIR RESULTS USING EOSUPPORT SOFTWARE

This article introduces a computer program EOSupport (Experimental Optimization Support) written by the author which helps to design experiments, process experimental data, calculate regression coefficients and compute statistical criterias for testing experimental results. The article also contains results of two successfully solved with EOSupport typical optimization problems.

ВОЛКОВ Виталий Витальевич (р.1984), аспирант (научный рук. Р.В.Нейдорф) первого года обучения кафедры «ПОВТ и АС» ДГТУ. Окончил ДГТУ по специальности «Программное обеспечение ВТ и АС» в 2007 году.

Область научных интересов: теория управления, планирование эксперимента, языки программирования, базы данных.

Автор 5 научных статей.